AU 1812

INPUT SET: S6822.raw

```
 1                              SEQUENCE LISTING
 2
 3   (1)    General Information:                 see p. 5 → 25
 4
 5        (i) APPLICANT: Breece, Tim
 6                      Hayenga, Kirk
 7                      Rinderknecht, Ernst
 8                      Vandlen, Richard
 9                      Yansura, Daniel
10
11        (ii) TITLE OF INVENTION: PROCESS FOR PRODUCING RELAXIN
12
13      (iii) NUMBER OF SEQUENCES: 40
14
15       (iv) CORRESPONDENCE ADDRESS:
16             (A) ADDRESSEE: Mr. Walter H. Dreger
17             (B) STREET: 4 Embarcadero Center, Suite 3400
18             (C) CITY: San Francisco
19             (D) STATE: California
20             (E) COUNTRY: USA
21            .(F) ZIP: 94111
22
23        (v) COMPUTER READABLE FORM:
24             (A) MEDIUM TYPE: Floppy disk
25             (B) COMPUTER: IBM PC compatible
26             (C) OPERATING SYSTEM: PC-DOS/MS-DOS
27             (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
28
29       (vi) CURRENT APPLICATION DATA:
30             (A) APPLICATION NUMBER: US
31             (B) FILING DATE: 21-JUN-1993
32             (C) CLASSIFICATION:
33
34     (viii) ATTORNEY/AGENT INFORMATION:
35             (A) NAME: Dreger, Walter H.
36             (B) REGISTRATION NUMBER: 24,190
37             (C) REFERENCE/DOCKET NUMBER: A-58117/WHD
38
39       (ix) TELECOMMUNICATION INFORMATION:
40             (A) TELEPHONE: (415) 781-1989
41            . (B) TELEFAX: (415) 398-3249
42
43
44   (2) INFORMATION FOR SEQ ID NO:1:
45
46        (i) SEQUENCE CHARACTERISTICS:
47             (A) LENGTH: 11 amino acids
48             (B) TYPE: amino acid
49             (C) STRANDEDNESS: single
50             (D) TOPOLOGY: linear
51
```

RAW SEQUENCE LISTING
PATENT APPLICATION *US/08/080,354*

INPUT SET: S6822.raw

```
52          (ii) MOLECULE TYPE: peptide
53
54
55
56          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:
57
58          Met Lys Lys Asn Ile Ala Phe Leu Leu Lys Arg
59          1               5                   10
60
61     (2) INFORMATION FOR SEQ ID NO:2:
62
63          (i) SEQUENCE CHARACTERISTICS:
64              (A) LENGTH: 29 amino acids
65              (B) TYPE: amino acid
66              (C) STRANDEDNESS: single
67              (D) TOPOLOGY: linear
68
69          (ii) MOLECULE TYPE: peptide
70
71
72
73          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:
74
75          Asp Ser Trp Met Glu Glu Val Ile Lys Leu Cys Gly Arg Glu Leu Val
76          1               5                   10                  15
77
78          Arg Ala Gln Ile Ala Ile Cys Gly Met Ser Thr Trp Ser
79                      20                  25
80
81     (2) INFORMATION FOR SEQ ID NO:3:
82
83          (i) SEQUENCE CHARACTERISTICS:
84              (A) LENGTH: 13 amino acids
85              (B) TYPE: amino acid
86              (C) STRANDEDNESS: single
87              (D) TOPOLOGY: linear
88
89          (ii) MOLECULE TYPE: peptide
90
91
92
93          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:
94
95          Lys Arg Lys Pro Thr Gly Tyr Gly Ser Arg Lys Lys Arg
96          1               5                   10
97
98     (2) INFORMATION FOR SEQ ID NO:4:
99
100         (i) SEQUENCE CHARACTERISTICS:
101             (A) LENGTH: 24 amino acids
102             (B) TYPE: amino acid
```

```
103                 (C)  STRANDEDNESS: single
104                 (D)  TOPOLOGY: linear
105
106          (ii)  MOLECULE TYPE: peptide
107
108
109
110          (xi)  SEQUENCE DESCRIPTION: SEQ ID NO:4:
111
112          Gln Leu Tyr Ser Ala Leu Ala Asn Lys Cys Cys His Val Gly Cys Thr
113          1               5                   10                  15
114
115          Lys Arg Ser Leu Ala Arg Phe Cys
116                       20
117
118   (2)  INFORMATION FOR SEQ ID NO:5:
119
120          (i)  SEQUENCE CHARACTERISTICS:
121               (A)  LENGTH: 13 amino acids
122               (B)  TYPE: amino acid
123               (C)  STRANDEDNESS: single
124               (D)  TOPOLOGY: linear
125
126          (ii)  MOLECULE TYPE: peptide
127
128
129
130          (xi)  SEQUENCE DESCRIPTION: SEQ ID NO:5:
131
132          Lys Arg Lys Pro Thr Gly Tyr Gly Ser Arg Lys Lys Arg
133          1               5                   10
134
135   (2)  INFORMATION FOR SEQ ID NO:6:
136
137          (i)  SEQUENCE CHARACTERISTICS:
138               (A)  LENGTH: 13 amino acids
139               (B)  TYPE: amino acid
140               (C)  STRANDEDNESS: single
141               (D)  TOPOLOGY: linear
142
143          (ii)  MOLECULE TYPE: peptide
144
145
146
147          (xi)  SEQUENCE DESCRIPTION: SEQ ID NO:6:
148
149          Asp Lys Lys Arg Thr Gly Tyr Gly Ser Arg Arg Arg Lys
150          1               5                   10
151
152   (2)  INFORMATION FOR SEQ ID NO:7:
153
```

```
154          (i) SEQUENCE CHARACTERISTICS:
155              (A) LENGTH: 13 amino acids
156              (B) TYPE: amino acid
157              (C) STRANDEDNESS: single
158              (D) TOPOLOGY: linear
159
160         (ii) MOLECULE TYPE: peptide
161
162
163
164         (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:
165
166          Asp Lys Lys Arg Thr Gly Tyr Gly Ser Arg Lys Lys Arg
167          1               5                   10
168
169  (2) INFORMATION FOR SEQ ID NO:8:
170
171          (i) SEQUENCE CHARACTERISTICS:
172              (A) LENGTH: 13 amino acids
173              (B) TYPE: amino acid
174              (C) STRANDEDNESS: single
175              (D) TOPOLOGY: linear
176
177         (ii) MOLECULE TYPE: peptide
178
179
180
181         (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:
182
183          Lys Arg Lys Pro Thr Gly Tyr Gly Ser Arg Arg Arg Lys
184          1               5                   10
185
186  (2) INFORMATION FOR SEQ ID NO:9:
187
188          (i) SEQUENCE CHARACTERISTICS:
189              (A) LENGTH: 231 base pairs
190              (B) TYPE: nucleic acid
191              (C) STRANDEDNESS: single
192              (D) TOPOLOGY: linear
193
194         (ii) MOLECULE TYPE: cDNA
195
196
197         (ix) FEATURE:
198              (A) NAME/KEY: CDS
199              (B) LOCATION: 1..231
200
201
202         (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:
203
204  ATG AAA AAG AAT ATC GCA TTT CTT CTT AAA CGG GAC TCA TGG
```

INPUT SET: S6822.raw

```
-->   205   ATG GAG           48
      206   Met Lys Lys Asn Ile Ala Phe Leu Leu Lys Arg Asp Ser Trp Met Glu
      207    1               5                   10                  15
      208
      209   GAA GTT ATT AAA TTA TGC GGC CGC GAA TTA GTT CGC GCG
-->   210   CAG ATT GCC           96
      211   Glu Val Ile Lys Leu Cys Gly Arg Glu Leu Val Arg Ala Gln Ile Ala
      212               20                  25                  30
      213
      214   ATT TGC GGC ATG AGC ACC TGG AGC AAA AGG AAA CCC ACT
-->   215   GGT TAT GGT           144
      216   Ile Cys Gly Met Ser Thr Trp Ser Lys Arg Lys Pro Thr Gly Tyr Gly
      217               35                  40                  45
      218
      219   TCT CGA AAA AAG AGA CAA CTC TAC AGT GCA TTG GCT AAT AAA
-->   220   TGT TGC           192
      221   Ser Arg Lys Lys Arg Gln Leu Tyr Ser Ala Leu Ala Asn Lys Cys Cys
      222       50                  55                  60
      223
      224   CAT GTT GGT TGT ACC AAA AGA TCT CTT GCT AGA TTT TGC
      225           231
      226   His Val Gly Cys Thr Lys Arg Ser Leu Ala Arg Phe Cys
      227    65                  70                  75
      228
      229
-->   230   (2)  INFORMATION FOR SEQ ID NO:34
      231
      232          (i)  SEQUENCE CHARACTERISTICS:
      233               (A)  LENGTH: 77 amino acids
      234               (B)  TYPE: amino acid
      235               (D)  TOPOLOGY: linear
      236
      237          (ii)  MOLECULE TYPE: protein
      238
      239          (xi)  SEQUENCE DESCRIPTION: SEQ ID NO:34:
      240
      241   Met Lys Lys Asn Ile Ala Phe Leu Leu Lys Arg Asp Ser Trp Met Glu
      242    1               5                   10                  15
      243
      244   Glu Val Ile Lys Leu Cys Gly Arg Glu Leu Val Arg Ala Gln Ile Ala
      245               20                  25                  30
      246
      247   Ile Cys Gly Met Ser Thr Trp Ser Lys Arg Lys Pro Thr Gly Tyr Gly
      248               35                  40                  45
      249
      250   Ser Arg Lys Lys Arg Gln Leu Tyr Ser Ala Leu Ala Asn Lys Cys Cys
      251       50                  55                  60
      252
      253   His Val Gly Cys Thr Lys Arg Ser Leu Ala Arg Phe Cys
      254    65                  70                  75
      255
```

[Handwritten annotations in right margin:] The nucleic number must be on the same line as each nucleic line. It appears that your margins were affected by importing the Patent In file into a word processor. Please widen margins.

✱ All nucleic sequences in this listing need to have the nucleic numbers wrap back up to the correct line.

[Handwritten note near line 230:] 10?

[Handwritten note near line 234:] What SEQ ID NO is this?

RAW SEQUENCE LISTING
PATENT APPLICATION *US/08/080,354*

INPUT SET: S6822.raw

```
 -->    256   (2) INFORMATION FOR SEQ ID NO:10:
        257
        258        (i) SEQUENCE CHARACTERISTICS:
 -->    259            (A) LENGTH: 593 base pairs
        260            (B) TYPE: nucleic acid
        261            (C) STRANDEDNESS: single
        262            (D) TOPOLOGY: linear
        263
        264       (ii) MOLECULE TYPE: cDNA
        265
        266
        267       (ix) FEATURE:
        268            (A) NAME/KEY: CDS
        269            (B) LOCATION: 431..586
        270
        271
        272       (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:
        273
        274   GAATTCAACT TCTCCATACT TTGGATAAGG AAATACAGAC
 -->    275   ATGAAAAATC TCATTGCTGA       60
        276
        277   GTTGTTATTT AAGCTTGCCC AAAAAGAAGA AGAGTCGAAA
 -->    278   GAACTGTGTG CGCAGGTAGA      120
        279
        280   AGCTTTGGAG ATTATCGTCA CTGCAATGCT TCGCAATATG
 -->    281   GCGCAAAATG ACCAACAGCG      180
        282
        283   GTTGATTGAT CAGGTAGAGG GGGCGCTGTA CGAGGTAAAG
 -->    284   CCCGATGCCA GCATTCCTGA      240
        285
        286   CGACGATACG GAGCTGCTGC GCGATTACGT AAAGAAGTTA
 -->    287   TTGAAGCATC CTCGTCAGTA      300
        288
        289   AAAAGTTAAT CTTTTCAACA GCTGTCATAA AGTTGTCACG
 -->    290   GCCGAGACTT ATAGTCGCTT      360
        291
        292   TGTTTTTATT TTTTAATGTA TTTGTACGCA AGTTCACGTA
 -->    293   AAAAGGGTAT CTAGAGGTTG      420
        294
        295   AGGTGATTTT ATG AAA AAG AAT ATC GCA TTT CTT CTT GCA TCT
 -->    296   ATG TTC        469
        297              Met Lys Lys Asn Ile Ala Phe Leu Leu Ala Ser Met Phe
        298                1               5                    10
        299
        300   GTT TTT TCT ATT GCT ACA AAT GCC TAT GCA GAC TCA TGG ATG
 -->    301   GAG GAA        517
        302   Val Phe Ser Ile Ala Thr Asn Ala Tyr Ala Asp Ser Trp Met Glu Glu
        303            15               20                   25
        304
        305   GTT ATT AAA TTA TGC GGC CGC GAA TTG GTA CGC GCG CAA ATA
 -->    306   GCG ATA        565
```

[handwritten annotation:] Please correct the sequence of SEQ ID NO's. The sequence on the previous page seems out of place

```
307    Val Ile Lys Leu Cys Gly Arg Glu Leu Val Arg Ala Gln Ile Ala Ile
308     30              35                  40                  45
309
310    TGC GGT ATG AGT ACA TGG AGT TGAAGAA
311     593
312    Cys Gly Met Ser Thr Trp Ser
313             50
314
315
316    (2) INFORMATION FOR SEQ ID NO:37:
317
318         (i) SEQUENCE CHARACTERISTICS:
319             (A) LENGTH: 52 amino acids
320             (B) TYPE: amino acid
321             (D) TOPOLOGY: linear
322
323         (ii) MOLECULE TYPE: protein
324
325         (xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:
326
327    Met Lys Lys Asn Ile Ala Phe Leu Leu Ala Ser Met Phe Val Phe Ser
328     1               5                  10                  15
329
330    Ile Ala Thr Asn Ala Tyr Ala Asp Ser Trp Met Glu Glu Val Ile Lys
331                 20                  25                  30
332
333    Leu Cys Gly Arg Glu Leu Val Arg Ala Gln Ile Ala Ile Cys Gly Met
334             35                  40                  45
335
336    Ser Thr Trp Ser
337         50
338
339    (2) INFORMATION FOR SEQ ID NO:11:
340
341         (i) SEQUENCE CHARACTERISTICS:
342             (A) LENGTH: 1500 base pairs
343             (B) TYPE: nucleic acid
344             (C) STRANDEDNESS: single
345             (D) TOPOLOGY: linear
346
347         (ii) MOLECULE TYPE: cDNA
348
349
350         (ix) FEATURE:
351             (A) NAME/KEY: CDS
352             (B) LOCATION: 438..1238
353
354
355         (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:
356
357    GAATTCAACT TCTCCATACT TTGGATAAGG AAATACAGAC
```

INPUT SET: S6822.raw

```
-->   358   ATGAAAAATC TCATTGCTGA        60
      359
      360   GTTGTTATTT AAGCTTGCCC AAAAAGAAGA AGAGTCGAAA
-->   361   GAACTGTGTG CGCAGGTAGA       120
      362
      363   AGCTTTGGAG ATTATCGTCA CTGCAATGCT TCGCAATATG
-->   364   GCGCAAAATG ACCAACAGCG       180
      365
      366   GTTGATTCAT CAGGTAGAGG GGGCGCTGTA CGAGGTAAAG
-->   367   CCCGATGCCA GCATTCCTGA       240
      368
      369   CGACGATACG GAGCTGCTGC GCGATTACGT AAAGAAGTTA
-->   370   TTGAAGCATC CTCGTCAGTA       300
      371
      372   AAAAGTTAAT CTTTTCAACA GCTGTCATAA AGTTGTCACG
-->   373   GCCGAGACTT ATAGTCGCTT       360
      374
      375   TGTTTTTATT TTTTAATGTA TTTGTAACTA GTACGCAAGT
-->   376   TCACGTAAAA AGGGTATCTA       420
      377
      378   GAGGTTGAGG TGATTTT ATG AAA AAG AAT ATC GCA TTT CTT
-->   379   CTT GCA TCT        470
      380                   Met Lys Lys Asn Ile Ala Phe Leu Leu Ala Ser
      381                    1               5                  10
      382
      383   ATG TTC GTT TTT TCT ATT GCT ACA AAT GCC TAT GCA TCA GGC
-->   384   ACT ACA       518
      385   Met Phe Val Phe Ser Ile Ala Thr Asn Ala Tyr Ala Ser Gly Thr Thr
      386              15               20                  25
      387
      388   AAT ACT GTG GCA GCA TAT AAT TTA ACT TGG AAA TCA ACT AAT
-->   389   TTC AAG       566
      390   Asn Thr Val Ala Ala Tyr Asn Leu Thr Trp Lys Ser Thr Asn Phe Lys
      391            30               35                  40
      392
      393   ACA ATT TTG GAG TGG GAA CCC AAA CCC GTC AAT CAA GTC TAC
-->   394   ACT GTT       614
      395   Thr Ile Leu Glu Trp Glu Pro Lys Pro Val Asn Gln Val Tyr Thr Val
      396        45               50                  55
      397
      398   CAA ATA AGC ACT AAG TCA GGA GAT TGG AAA AGC AAA TGC
-->   399   TTT TAC ACA        662
      400   Gln Ile Ser Thr Lys Ser Gly Asp Trp Lys Ser Lys Cys Phe Tyr Thr
      401   60               65                  70                  75
      402
      403   ACA GAC ACA GAG TGT GAC CTC ACC GAC GAG ATT GTG AAG
-->   404   GAT GTG AAG        710
      405   Thr Asp Thr Glu Cys Asp Leu Thr Asp Glu Ile Val Lys Asp Val Lys
      406              80               85                  90
      407
      408   CAG ACG TAC TTG GCA CGG GTC TTC TCC TAC CCG GCA GGG AAT
```

```
-->   409   GTG GAG          758
      410   Gln Thr Tyr Leu Ala Arg Val Phe Ser Tyr Pro Ala Gly Asn Val Glu
      411            95                  100              105
      412
      413   AGC ACC GGT TCT GCT GGG GAG CCT CTG TAT GAG AAC TCC CCA
-->   414   GAG TTC          806
      415   Ser Thr Gly Ser Ala Gly Glu Pro Leu Tyr Glu Asn Ser Pro Glu Phe
      416            110                 115              120
      417
      418   ACA CCT TAC CTG GAG ACA AAC CTC GGA CAG CCA ACA ATT CAG
-->   419   AGT TTT          854
      420   Thr Pro Tyr Leu Glu Thr Asn Leu Gly Gln Pro Thr Ile Gln Ser Phe
      421            125                 130              135
      422
      423   GAA CAG GTG GGA ACA AAA GTG AAT GTG ACC GTA GAA GAT
-->   424   GAA CGG ACT          902
      425   Glu Gln Val Gly Thr Lys Val Asn Val Thr Val Glu Asp Glu Arg Thr
      426   140                 145              150                  155
      427
      428   TTA GTC AGA AGG AAC AAC ACT TTC CTA AGC CTC CGG GAT GTT
-->   429   TTT GGC          950
      430   Leu Val Arg Arg Asn Asn Thr Phe Leu Ser Leu Arg Asp Val Phe Gly
      431                160                 165                  170
      432
      433   AAG GAC TTA ATT TAT ACA CTT TAT TAT TGG AAA TCT TCA AGT
-->   434   TCA GGA          998
      435   Lys Asp Leu Ile Tyr Thr Leu Tyr Tyr Trp Lys Ser Ser Ser Ser Gly
      436                175                 180              185
      437
      438   AAG AAA ACA GCC AAA ACA AAC ACT AAT GAG TTT TTG ATT GAT
-->   439   GTG GAT          1046
      440   Lys Lys Thr Ala Lys Thr Asn Thr Asn Glu Phe Leu Ile Asp Val Asp
      441            190                 195              200
      442
      443   AAA GGA GAA AAC TAC TGT TTC AGT GTT CAA GCA GTG ATT CCC
-->   444   TCC CGA          1094
      445   Lys Gly Glu Asn Tyr Cys Phe Ser Val Gln Ala Val Ile Pro Ser Arg
      446            205                 210              215
      447
      448   ACA GTT AAC CGG AAG AGT ACA GAC AGC CCG GTA GAG TGT
-->   449   ATG GGC CAG          1142
      450   Thr Val Asn Arg Lys Ser Thr Asp Ser Pro Val Glu Cys Met Gly Gln
      451   220                 225              230                  235
      452
      453   GAG AAA GGC CAA TTC AGA GAA ATA TTC TAC ATC ATT GGA GCT
-->   454   GTG GTA          1190
      455   Glu Lys Gly Gln Phe Arg Glu Ile Phe Tyr Ile Ile Gly Ala Val Val
      456            240                 245              250
      457
      458   TTT GTG GTC ATC ATC CTT GTC ATC ATC CTG GCT ATA TCT CTA
-->   459   CAC TAAAATTCTC          1245
```

INPUT SET: S6822.raw

```
460   Phe Val Val Ile Ile Leu Val Ile Ile Leu Ala Ile Ser Leu His
461           255                 260                 265
462
463   ATGTTTGACA GCTTATCATC GATAAGCTTT AATGCGGTAG
--> 464   TTTATCACAG TTAAATTGCT     1305
465
466   AACGCAGTCA GGCACCGTGT ATGAAATCTA ACAATGCGCT
--> 467   CATCGTCATC CTCGGCACCG     1365
468
469   TCACCCTGGA TGCTGTAGGC ATAGGCTTGG TTATGCCGGT
--> 470   ACTGCCGGGC CTCTTGCGGG     1425
471
472   ATATCGTCCA TTCCGACAGC ATCGCCAGTC ACTATGGCGT
--> 473   GCTCCTAGCG CTATATGCGT     1485
474
475   TGATGCAATT TCTAT                              1500
476
477
--> 478   (2) INFORMATION FOR SEQ ID NO:38:
479
480           (i) SEQUENCE CHARACTERISTICS:
481                   (A) LENGTH: 266 amino acids
482                   (B) TYPE: amino acid
483                   (D) TOPOLOGY: linear
484
485           (ii) MOLECULE TYPE: protein
486
487           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:
488
489   Met Lys Lys Asn Ile Ala Phe Leu Leu Ala Ser Met Phe Val Phe Ser
490     1               5                 10                 15
491
492   Ile Ala Thr Asn Ala Tyr Ala Ser Gly Thr Thr Asn Thr Val Ala Ala
493               20                25                 30
494
495   Tyr Asn Leu Thr Trp Lys Ser Thr Asn Phe Lys Thr Ile Leu Glu Trp
496           35                40                 45
497
498   Glu Pro Lys Pro Val Asn Gln Val Tyr Thr Val Gln Ile Ser Thr Lys
499         50                55                 60
500
501   Ser Gly Asp Trp Lys Ser Lys Cys Phe Tyr Thr Thr Asp Thr Glu Cys
502     65                70                75                 80
503
504   Asp Leu Thr Asp Glu Ile Val Lys Asp Val Lys Gln Thr Tyr Leu Ala
505               85                90                 95
506
507   Arg Val Phe Ser Tyr Pro Ala Gly Asn Val Glu Ser Thr Gly Ser Ala
508           100               105               110
509
510   Gly Glu Pro Leu Tyr Glu Asn Ser Pro Glu Phe Thr Pro Tyr Leu Glu
```

```
511              115                   120                   125
512
513     Thr Asn Leu Gly Gln Pro Thr Ile Gln Ser Phe Glu Gln Val Gly Thr
514          130                   135                   140
515
516     Lys Val Asn Val Thr Val Glu Asp Glu Arg Thr Leu Val Arg Arg Asn
517     145                   150                   155                   160
518
519     Asn Thr Phe Leu Ser Leu Arg Asp Val Phe Gly Lys Asp Leu Ile Tyr
520                       165                   170                   175
521
522     Thr Leu Tyr Tyr Trp Lys Ser Ser Ser Gly Lys Lys Thr Ala Lys
523                   180                   185                   190
524
525     Thr Asn Thr Asn Glu Phe Leu Ile Asp Val Asp Lys Gly Glu Asn Tyr
526              195                   200                   205
527
528     Cys Phe Ser Val Gln Ala Val Ile Pro Ser Arg Thr Val Asn Arg Lys
529          210                   215                   220
530
531     Ser Thr Asp Ser Pro Val Glu Cys Met Gly Gln Glu Lys Gly Gln Phe
532     225                   230                   235                   240
533
534     Arg Glu Ile Phe Tyr Ile Ile Gly Ala Val Val Phe Val Val Ile Ile
535                   245                   250                   255
536
537     Leu Val Ile Ile Leu Ala Ile Ser Leu His
538                   260                   265
539
540
541     (2) INFORMATION FOR SEQ ID NO:12:
542
543         (i) SEQUENCE CHARACTERISTICS:
544             (A) LENGTH: 11 amino acids
545             (B) TYPE: amino acid
546             (C) STRANDEDNESS: single
547             (D) TOPOLOGY: linear
548
549         (ii) MOLECULE TYPE: peptide
550
551
552
553         (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:
554
555         Met Lys Lys Asn Ile Ala Phe Leu Leu Arg Lys
556         1               5                   10
557
558     (2) INFORMATION FOR SEQ ID NO:13:
559
560         (i) SEQUENCE CHARACTERISTICS:
561             (A) LENGTH: 11 amino acids
```

PAGE: 12

RAW SEQUENCE LISTING
PATENT APPLICATION *US/08/080,354*

DATE: 10/27/93
TIME: 15:19:14

INPUT SET: S6822.raw

```
562                    (B)  TYPE: amino acid
563                    (C)  STRANDEDNESS: single
564                    (D)  TOPOLOGY: linear
565
566          (ii)  MOLECULE TYPE: peptide
567
568
569
570          (xi)  SEQUENCE DESCRIPTION: SEQ ID NO:13:
571
572          Met Lys Lys Asn Ile Ala Phe Leu Leu Arg Arg
573           1               5                  10
574
575  (2)  INFORMATION FOR SEQ ID NO:14:
576
577          (i)  SEQUENCE CHARACTERISTICS:
578                    (A)  LENGTH: 11 amino acids
579                    (B)  TYPE: amino acid
580                    (C)  STRANDEDNESS: single
581                    (D)  TOPOLOGY: linear
582
583          (ii)  MOLECULE TYPE: peptide
584
585
586
587          (xi)  SEQUENCE DESCRIPTION: SEQ ID NO:14:
588
589          Met Lys Lys Asn Ile Ala Phe Leu Leu Lys Lys
590           1               5                  10
591
592  (2)  INFORMATION FOR SEQ ID NO:15:
593
594          (i)  SEQUENCE CHARACTERISTICS:
595                    (A)  LENGTH: 42 base pairs
596                    (B)  TYPE: nucleic acid
597                    (C)  STRANDEDNESS: both
598                    (D)  TOPOLOGY: linear
599
600          (ii)  MOLECULE TYPE: cDNA
601
602
603          (ix)  FEATURE:
604                    (A)  NAME/KEY: misc_feature
605                    (B)  LOCATION: 5..42
606                   -(D)  OTHER INFORMATION: /note= "Complementary
607                         double-stranded binding between bases 5 and 42 to
608                         SEQ ID NO:16. "
609
610
611          (xi)  SEQUENCE DESCRIPTION: SEQ ID NO:15:
612
```

```
613    CTAGAATTAT GAAAAAGAAT ATCGCATTTC TTCTTAAACG GG
614                     42
615
616    (2) INFORMATION FOR SEQ ID NO:16:
617
618         (i)  SEQUENCE CHARACTERISTICS:
619              (A) LENGTH: 41 base pairs
620              (B) TYPE: nucleic acid
621              (C) STRANDEDNESS: both
622              (D) TOPOLOGY: linear
623
624         (ii) MOLECULE TYPE: cDNA
625
626
627         (ix) FEATURE:
628              (A) NAME/KEY: misc_feature
629              (B) LOCATION: 4..41
630              (D) OTHER INFORMATION: /note= "Complementary
631                   double-stranded binding between bases 4 and 41 to
632                   SEQ ID NO:15."
633
634
635         (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:
636
637    AGTCCCGTTT AAGAAGAAAT GCGATATTCT TTTTCATAAT T
638                     41
639
640    (2) INFORMATION FOR SEQ ID NO:17:
641
642         (i)  SEQUENCE CHARACTERISTICS:
643              (A) LENGTH: 42 base pairs
644              (B) TYPE: nucleic acid
645              (C) STRANDEDNESS: both
646              (D) TOPOLOGY: linear
647
648         (ii) MOLECULE TYPE: cDNA
649
650
651         (ix) FEATURE:
652              (A) NAME/KEY: misc_feature
653              (B) LOCATION: 5..42
654              (D) OTHER INFORMATION: /note= "Complementary
655                   double-stranded binding between bases 5 and 42 to
656                   SEQ ID NO:18."
657
658         (ix) FEATURE:
659              (A) NAME/KEY: CDS
660              (B) LOCATION: 9..41
661
662
663         (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:
```

--> 619

--> 643

PAGE: 14

RAW SEQUENCE LISTING
PATENT APPLICATION US/08/080,354

DATE: 10/27/93
TIME: 15:19:28

INPUT SET: S6822.raw

```
664
665     CTAGAATT ATG TTC CCA GCT ATG CCT CTA TCT AGT AAA CGG G
666                   42
667              Met Phe Pro Ala Met Pro Leu Ser Ser Lys Arg
668               1           5                   10
669
670
671     (2) INFORMATION FOR SEQ ID NO:19:
672
673              (i) SEQUENCE CHARACTERISTICS:
674                      (A) LENGTH: 11 amino acids
675                      (B) TYPE: amino acid
676                      (D) TOPOLOGY: linear
677
678          (ii) MOLECULE TYPE: protein
679
680          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:
681
682     Met Phe Pro Ala Met Pro Leu Ser Ser Lys Arg
683      1           5                   10
684
685     (2) INFORMATION FOR SEQ ID NO:18:
686
687              (i) SEQUENCE CHARACTERISTICS:
688                      (A) LENGTH: 41 base pairs
689                      (B) TYPE: nucleic acid
690                      (C) STRANDEDNESS: both
691                      (D) TOPOLOGY: linear
692
693          (ii) MOLECULE TYPE: cDNA
694
695
696          (ix) FEATURE:
697                      (A) NAME/KEY: misc_feature
698                      (B) LOCATION: 4..41
699                      (D) OTHER INFORMATION: /note= "Complementary
700                          double-stranded binding between bases 4 and 41 to
701                          SEQ ID NO:17."
702
703
704          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:
705
706     AGTCCCGTTT ACTAGATAGA GGCATAGCTG GGAACATAAT T
707                   41
708
709     (2) INFORMATION FOR SEQ ID NO:20:
710
711              (i) SEQUENCE CHARACTERISTICS:
712                      (A) LENGTH: 64 base pairs
713                      (B) TYPE: nucleic acid
714                      (C) STRANDEDNESS: both
```

Shouldn't this go BEFORE SEQ 19?

RAW SEQUENCE LISTING
PATENT APPLICATION *US/08/080,354*

INPUT SET: S6822.raw

```
715                    (D) TOPOLOGY: linear
716
717         (ii) MOLECULE TYPE: cDNA
718
719
720         (ix) FEATURE:
721                    (A) NAME/KEY: CDS
722                    (B) LOCATION: 5..64
723
724         (ix) FEATURE:
725                    (A) NAME/KEY: misc_feature
726                    (B) LOCATION: 5..64
727                    (D) OTHER INFORMATION: /note= "Complementary
728                        double-stranded binding between bases 5 and 64 to
729                        SEQ ID NO:21."
730
731
732         (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:
733
734  CGCG CAG ATT GCC ATT TGC GGC ATG AGC ACC TGG AGC AAA
735  AGG AAA CCC          49
736        Gln Ile Ala Ile Cys Gly Met Ser Thr Trp Ser Lys Arg Lys Pro
737         1               5                  10                  15
738
739  ACT GGT TAT GGT TCT                                            64
740  Thr Gly Tyr Gly Ser
741                    20
742
743
744    (2) INFORMATION FOR SEQ ID NO:22:
745
746            (i) SEQUENCE CHARACTERISTICS:
747                    (A) LENGTH: 20 amino acids
748                    (B) TYPE: amino acid
749                    (D) TOPOLOGY: linear
750
751           (ii) MOLECULE TYPE: protein
752
753           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:
754
755  Gln Ile Ala Ile Cys Gly Met Ser Thr Trp Ser Lys Arg Lys Pro Thr
756   1               5                  10                  15
757
758  Gly Tyr Gly Ser
759                    20
760
761    (2) INFORMATION FOR SEQ ID NO:21:
762
763            (i) SEQUENCE CHARACTERISTICS:
764                    (A) LENGTH: 62 base pairs
765                    (B) TYPE: nucleic acid
```

(handwritten annotation) Place this AFTER SEQ 21.

```
        766              (C) STRANDEDNESS: both
        767              (D) TOPOLOGY: linear
        768
        769         (ii) MOLECULE TYPE: cDNA
        770
        771
        772         (ix) FEATURE:
        773              (A) NAME/KEY: misc_feature
        774              (B) LOCATION: 3..62
        775              (D) OTHER INFORMATION: /note= "Complementary
        776                  double-stranded binding between bases 3 and 62 to
        777                  SEQ ID NO:20."
        778
        779
        780         (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:
        781
        782    CGAGAACCAT AACCAGTGGG TTTCCTTTTG CTCCAGGTGC
-->     783    TCATGCCGCA AATGGCAATC        60
        784
        785    TG                                                      62
        786
-->     787    (2) INFORMATION FOR SEQ ID NO:23:
        788
        789         (i) SEQUENCE CHARACTERISTICS:
-->     790              (A) LENGTH: 50 base pairs
        791              (B) TYPE: nucleic acid
        792              (C) STRANDEDNESS: both
        793              (D) TOPOLOGY: linear
        794
        795         (ii) MOLECULE TYPE: cDNA
        796
        797
        798         (ix) FEATURE:
        799              (A) NAME/KEY: misc_feature
        800              (B) LOCATION: 5..50
        801              (D) OTHER INFORMATION: /note= "Complementary
        802                  double-stranded binding between bases 5 and 50
        803                  with SEQ ID NO:24."
        804
        805
        806         (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:
        807
        808    GGCCACTCTG TGCGGTGCTG AACTGGTTGA CGCTCTGCAG
-->     809    TTTGTTTGCG                  50
        810
        811    (2) INFORMATION FOR SEQ ID NO:24:
        812
        813         (i) SEQUENCE CHARACTERISTICS:
-->     814              (A) LENGTH: 50 base pairs
        815              (B) TYPE: nucleic acid
        816              (C) STRANDEDNESS: both
```

PAGE: 17

RAW SEQUENCE LISTING
PATENT APPLICATION *US/08/080,354*

DATE: 10/27/93
TIME: 15:19:48

INPUT SET: S6822.raw

```
817                   (D) TOPOLOGY: linear
818
819          (ii) MOLECULE TYPE: cDNA
820
821
822          (ix) FEATURE:
823                   (A) NAME/KEY: misc_feature
824                   (B) LOCATION: 6..50
825                   (D) OTHER INFORMATION: /note= "Complementary
826                        double-stranded binding between bases 6 and 50
827                        with SEQ ID NO:23."
828
829
830          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:
831
832     GTCACCGCAA ACAAACTGCA GAGCGTCAAC CAGTTCAGCA
833     CCGCAAGAGT                    50
834
835     (2) INFORMATION FOR SEQ ID NO:25:
836
837          (i) SEQUENCE CHARACTERISTICS:
838                   (A) LENGTH: 55 base pairs
839                   (B) TYPE: nucleic acid
840                   (C) STRANDEDNESS: both
841                   (D) TOPOLOGY: linear
842
843          (ii) MOLECULE TYPE: cDNA
844
845
846          (ix) FEATURE:
847                   (A) NAME/KEY: misc_feature
848                   (B) LOCATION: 1..55
849                   (D) OTHER INFORMATION: /note= "Complementary
850                        double-stranded binding between bases 1 and 55
851                        with SEQ ID NO:26."
852
853
854          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:
855
856     GGTCCCGAAA CTCTGTGCGG TGCTGAACTG GTTGACGCTC
857     TGCAGTTTGT TTGCG                55
858
859     (2) INFORMATION FOR SEQ ID NO:26:
860
861          (i) SEQUENCE CHARACTERISTICS:
862                   (A) LENGTH: 64 base pairs
863                   (B) TYPE: nucleic acid
864                   (C) STRANDEDNESS: both
865                   (D) TOPOLOGY: linear
866
867          (ii) MOLECULE TYPE: cDNA
```

```
        868
        869
        870        (ix) FEATURE:
        871              (A) NAME/KEY: misc_feature
        872              (B) LOCATION: 6..64
        873              (D) OTHER INFORMATION: /note= "Complementary
        874                   double-stranded binding between bases 6 and 64
        875                   with SEQ ID NO:25."
        876
        877
        878        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:
        879
        880   GTCACCGCAA ACAAACTGCA GAGCGTCAAC CAGTTCAGCA
-->     881   CCGCACAGAG TTTCGGGACC          60
        882
        883   TGCA                                                          64
        884
        885   (2) INFORMATION FOR SEQ ID NO:27:
        886
        887        (i) SEQUENCE CHARACTERISTICS:
-->     888              (A) LENGTH: 84 base pairs
        889              (B) TYPE: nucleic acid
        890              (C) STRANDEDNESS: both
        891              (D) TOPOLOGY: linear
        892
        893        (ii) MOLECULE TYPE: cDNA
        894
        895
        896        (ix) FEATURE:
        897              (A) NAME/KEY: misc_feature
        898              (B) LOCATION: 5..84
        899              (D) OTHER INFORMATION: /note= "Complementary
        900                   double-stranded binding between bases 5 and 84
        901                   with SEQ ID NO:28."
        902
        903
        904        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:
        905
        906   CTAGAATTAT GATGATTACT CTGCGCAAAC TTCCTCTGGC
-->     907   GGTTGCCGTC GCAGCGGGCG          60
        908
        909   TAATGTCTGC TCAGGCCATG GCCA                                    84
        910
        911   (2) INFORMATION FOR SEQ ID NO:28:
        912
        913        (i) SEQUENCE CHARACTERISTICS:
-->     914              (A) LENGTH: 84 base pairs
        915              (B) TYPE: nucleic acid
        916              (C) STRANDEDNESS: both
        917              (D) TOPOLOGY: linear
        918
```

PAGE: 19

RAW SEQUENCE LISTING
PATENT APPLICATION *US/08/080,354*

DATE: 10/27/93
TIME: 15:20:02

INPUT SET: S6822.raw

```
919        (ii) MOLECULE TYPE: cDNA
920
921
922        (ix) FEATURE:
923             (A) NAME/KEY: misc_feature
924             (B) LOCATION: 5..84
925             (D) OTHER INFORMATION: /note= "Complementary
926                   double-stranded binding between bases 5 and 84
927                   with SEQ ID NO:27."
928
929
930        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:
931
932   GATCTGGCCA TGGCCTGAGC AGACATTACG CCCGCTGCGA
933   CGGCAACCGC CAGAGGAAGT          60
934
935   TTGCGCAGAG TAATCATCAT AATT                                    84
936
937   (2) INFORMATION FOR SEQ ID NO:29:
938
939        (i) SEQUENCE CHARACTERISTICS:
940             (A) LENGTH: 52 base pairs
941             (B) TYPE: nucleic acid
942             (C) STRANDEDNESS: both
943             (D) TOPOLOGY: linear
944
945        (ii) MOLECULE TYPE: cDNA
946
947
948        (ix) FEATURE:
949             (A) NAME/KEY: misc_feature
950             (B) LOCATION: 1..52
951             (D) OTHER INFORMATION: /note= "Complementary
952                   double-stranded binding between bases 1 and 52
953                   with SEQ ID NO:30."
954
955
956        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:
957
958   CAACTCTACA GTGCATTGGC TAATAAATGT TGCCATGTTG
959   GTTGTACCAA AA              52
960
961   (2) INFORMATION FOR SEQ ID NO:30:
962
963        (i) SEQUENCE CHARACTERISTICS:
964             (A) LENGTH: 60 base pairs
965             (B) TYPE: nucleic acid
966             (C) STRANDEDNESS: both
967             (D) TOPOLOGY: linear
968
969        (ii) MOLECULE TYPE: cDNA
```

```
970
971
972        (ix)  FEATURE:
973              (A)  NAME/KEY: misc_feature
974              (B)  LOCATION: 5..56
975              (D)  OTHER INFORMATION: /note= "Complementary
976                   double-stranded binding between bases 5 and 56
977                   with SEQ ID NO:29."
978
979
980        (xi)  SEQUENCE DESCRIPTION: SEQ ID NO:30:
981
982    GATCTTTTGG TACAACCAAC ATGGCAACAT TTATTAGCCA
-->  983    ATGCACTGTA GAGTTGTGCA          60
984
985
986    (2)  INFORMATION FOR SEQ ID NO:31:
987
988        (i)  SEQUENCE CHARACTERISTICS:
989              (A)  LENGTH: 13 amino acids
990              (B)  TYPE: amino acid
991              (C)  STRANDEDNESS: single
992              (D)  TOPOLOGY: linear
993
994        (ii)  MOLECULE TYPE: peptide
995
996
997
998        (xi)  SEQUENCE DESCRIPTION: SEQ ID NO:31:
999
1000       Asp Lys Lys Arg Thr Gly Tyr Gly Ser Arg Arg Arg Lys
1001        1               5                   10
1002
1003   (2)  INFORMATION FOR SEQ ID NO:32:
1004
1005       (i)  SEQUENCE CHARACTERISTICS:
1006              (A)  LENGTH: 13 amino acids
1007              (B)  TYPE: amino acid
1008              (C)  STRANDEDNESS: single
1009              (D)  TOPOLOGY: linear
1010
1011       (ii)  MOLECULE TYPE: peptide
1012
1013
1014
1015       (xi)  SEQUENCE DESCRIPTION: SEQ ID NO:32:
1016
1017       Asp Lys Lys Arg Thr Gly Tyr Gly Ser Arg Lys Lys Arg
1018        1               5                   10
1019
1020   (2)  INFORMATION FOR SEQ ID NO:33:
```

```
1021
1022          (i) SEQUENCE CHARACTERISTICS:
1023              (A) LENGTH: 13 amino acids
1024              (B) TYPE: amino acid
1025              (C) STRANDEDNESS: single
1026              (D) TOPOLOGY: linear
1027
1028         (ii) MOLECULE TYPE: peptide
1029
1030
1031
1032         (xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:
1033
1034              Lys Arg Lys Pro Thr Gly Tyr Gly Ser Arg Arg Arg Lys
1035              1               5                   10
1036
1037     (2) INFORMATION FOR SEQ ID NO:35:
1038
1039          (i) SEQUENCE CHARACTERISTICS:
1040              (A) LENGTH: 915 base pairs
1041              (B) TYPE: nucleic acid
1042              (C) STRANDEDNESS: single
1043              (D) TOPOLOGY: linear
1044
1045         (ii) MOLECULE TYPE: cDNA
1046
1047
1048         (ix) FEATURE:
1049              (A) NAME/KEY: CDS
1050              (B) LOCATION: 3..452
1051
1052
1053         (xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:
1054
1055     GC GGC CGC GAA TTA GTT CGC GCG CAG ATT GCC ATT TGC GGC
1056     ATG AGC           47
1057        Gly Arg Glu Leu Val Arg Ala Gln Ile Ala Ile Cys Gly Met Ser
1058        1               5                   10                  15
1059
1060     ACC TGG AGC AAA AGG TCT CTG AGC CAG GAA GAT GCT CCT
1061     CAG ACA CCT           95
1062     Thr Trp Ser Lys Arg Ser Leu Ser Gln Glu Asp Ala Pro Gln Thr Pro
1063                 20                  25                  30
1064
1065     AGA CCA GTG GCA GAA ATT GTG CCA TCC TTC ATC AAC AAA GAT
1066     ACA GAA           143
1067     Arg Pro Val Ala Glu Ile Val Pro Ser Phe Ile Asn Lys Asp Thr Glu
1068                 35                  40                  45
1069
1070     ACC ATA AAT ATG ATG TCA GAA TTT GTT GCT AAT TTG CCA CAG
1071     GAG CTG           191
```

(handwritten annotation, right margin:) ← IS SEQ 34 on p.5? Please place in order.

(handwritten arrows in left margin pointing to lines 1037, 1040, 1056, 1061, 1066, 1071)

INPUT SET: S6822.raw

```
1072    Thr Ile Asn Met Met Ser Glu Phe Val Ala Asn Leu Pro Gln Glu Leu
1073            50              55              60
1074
1075    AAG TTA ACC CTG TCT GAG ATG CAG CCA GCA TTA CCA CAG CTA
-->  1076    CAA CAA       239
1077    Lys Leu Thr Leu Ser Glu Met Gln Pro Ala Leu Pro Gln Leu Gln Gln
1078        65                  70              75
1079
1080    CAT GTA CCT GTA TTA AAA GAT TCC AGT CTT CTC TTT GAA GAA
-->  1081    TTT AAG       287
1082    His Val Pro Val Leu Lys Asp Ser Ser Leu Leu Phe Glu Glu Phe Lys
1083    80              85                  90                  95
1084
1085    AAA CTT ATT CGC AAT AGA CAA AGT GAA GCC GCA GAC AGC
-->  1086    AGT CCT TCA       335
1087    Lys Leu Ile Arg Asn Arg Gln Ser Glu Ala Ala Asp Ser Ser Pro Ser
1088            100             105             110
1089
1090    GAA TTA AAA TAC TTA GGC TTG GAT ACT CAT TCT CGA AAA AAG
-->  1091    AGA CAA       383
1092    Glu Leu Lys Tyr Leu Gly Leu Asp Thr His Ser Arg Lys Lys Arg Gln
1093            115             120             125
1094
1095    CTC TAC AGT GCA TTG GCT AAT AAA TGT TGC CAT GTT GGT TGT
-->  1096    ACC AAA       431
1097    Leu Tyr Ser Ala Leu Ala Asn Lys Cys Cys His Val Gly Cys Thr Lys
1098            130             135             140
1099
1100    AGA TCT CTT GCT AGA TTT TGC TGAGATGAAG CTAATTGTGC
-->  1101    ACATCTCGTA       482
1102    Arg Ser Leu Ala Arg Phe Cys
1103        145             150
1104
1105    TAATATTCAC ACATATTCTT AATGACATTT CACTGATGCT
-->  1106    TCTATCAGGT CAATTCTCAT       542
1107
1108    GTTTGACAGC TTATCATCGA TAAGCTTTAA TGCGGTAGTT
-->  1109    TATCACAGTT AAATTGCTAA       602
1110
1111    CGCAGTCAGG CACCGTGTAT GAAATCTAAC AATGCGCTCA
-->  1112    TCGTCATCCT CGGCACCGTC       662
1113
1114    ACCCTGGATG CTGTAGGCAT AGGCTTGGTT ATGCCGGTAC
-->  1115    TGCCGGGCCT CTTGCGGGAT       722
1116
1117    ATCGTCCATT CCGACAGCAT CGCCAGTCAC TATGGCGTGC
-->  1118    TGCTAGCGCT ATATGCGTTG       782
1119
1120    ATGCAATTTC TATGCGCACC CGTTCTCGGA GCACTGTCCG
-->  1121    ACCGCTTTGG CCGCCGCCCA       842
1122
```

```
      1123    GTCCTGCTCG CTTCGCTACT TGGAGCCACT ATCGACTACG
-->   1124    CGATCATGGC GACCACACCC        902
      1125
      1126    GTCCTGTGGA TCC                                           915
      1127
      1128
      1129    (2)  INFORMATION FOR SEQ ID NO:36:
      1130
      1131         (i)  SEQUENCE CHARACTERISTICS:
      1132              (A)  LENGTH: 150 amino acids
      1133              (B)  TYPE: amino acid
      1134              (D)  TOPOLOGY: linear
      1135
      1136         (ii)  MOLECULE TYPE: protein
      1137
      1138         (xi)  SEQUENCE DESCRIPTION: SEQ ID NO:36:
      1139
      1140    Gly Arg Glu Leu Val Arg Ala Gln Ile Ala Ile Cys Gly Met Ser Thr
      1141     1               5              10                  15
      1142
      1143    Trp Ser Lys Arg Ser Leu Ser Gln Glu Asp Ala Pro Gln Thr Pro Arg
      1144             20              25                  30
      1145
      1146    Pro Val Ala Glu Ile Val Pro Ser Phe Ile Asn Lys Asp Thr Glu Thr
      1147             35                  40                  45
      1148
      1149    Ile Asn Met Met Ser Glu Phe Val Ala Asn Leu Pro Gln Glu Leu Lys
      1150         50                  55                  60
      1151
      1152    Leu Thr Leu Ser Glu Met Gln Pro Ala Leu Pro Gln Leu Gln Gln His
      1153     65                  70                  75                  80
      1154
      1155    Val Pro Val Leu Lys Asp Ser Ser Leu Leu Phe Glu Glu Phe Lys Lys
      1156                 85                  90                  95
      1157
      1158    Leu Ile Arg Asn Arg Gln Ser Glu Ala Ala Asp Ser Ser Pro Ser Glu
      1159             100                 105                 110
      1160
      1161    Leu Lys Tyr Leu Gly Leu Asp Thr His Ser Arg Lys Lys Arg Gln Leu
      1162             115                 120                 125
      1163
      1164    Tyr Ser Ala Leu Ala Asn Lys Cys Cys His Val Gly Cys Thr Lys Arg
      1165         130                 135                 140
      1166
      1167    Ser Leu Ala Arg Phe Cys
      1168    145                 150
      1169
-->   1170    (2)  INFORMATION FOR SEQ ID NO:39:
      1171
      1172         (i)  SEQUENCE CHARACTERISTICS:
-->   1173              (A)  LENGTH: 301 base pairs
```

```
1174                    (B) TYPE: nucleic acid
1175                    (C) STRANDEDNESS: single
1176                    (D) TOPOLOGY: linear
1177
1178           (ii) MOLECULE TYPE: cDNA
1179
1180
1181           (ix) FEATURE:
1182                    (A) NAME/KEY: CDS
1183                    (B) LOCATION: 7..297
1184
1185
1186           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:
1187
1188      AAGCTT ATG AAA TCT AAC AAT GCG CTC ATC GTC ATC CTC GGC
-->  1189      ACC GTC          48
1190              Met Lys Ser Asn Asn Ala Leu Ile Val Ile Leu Gly Thr Val
1191               1               5                  10
1192
1193      ACC CTG GAT GCT GTA GGC ATA GGC TTG GTT ATG CCG GTA CTG
-->  1194      CCG GGC          96
1195      Thr Leu Asp Ala Val Gly Ile Gly Leu Val Met Pro Val Leu Pro Gly
1196       15               20               25               30
1197
1198      CTC TTG CGG GAT ATC GTC CAT TCC GAC AGC ATC GCC AGT CAC
-->  1199      TAT GGC          144
1200      Leu Leu Arg Asp Ile Val His Ser Asp Ser Ile Ala Ser His Tyr Gly
1201                   35               40               45
1202
1203      GTG CTG CTA GCG CTA TAT GCG TTG ATG CAA TTT CTA TGC GCA
-->  1204      CCC GTT          192
1205      Val Leu Leu Ala Leu Tyr Ala Leu Met Gln Phe Leu Cys Ala Pro Val
1206                   50               55               60
1207
1208      CTC GGA GCA CTG TCC GAC CGC TTT GGC CGC CGC CCA GTC CTG
-->  1209      CTC GCT          240
1210      Leu Gly Ala Leu Ser Asp Arg Phe Gly Arg Arg Pro Val Leu Leu Ala
1211                   65               70               75
1212
1213      TCG CTA CTT GGA GCC ACT ATC GAC TAC GCG ATC ATG GCG ACC
-->  1214      ACA CCC          288
1215      Ser Leu Leu Gly Ala Thr Ile Asp Tyr Ala Ile Met Ala Thr Thr Pro
1216           80               85               90
1217
1218      GTC CTG TGG ATCC                                              301
1219      Val Leu Trp
1220       95
1221
1222
1223      (2) INFORMATION FOR SEQ ID NO:40:
1224
```

```
1225            (i)  SEQUENCE CHARACTERISTICS:
1226                 (A)  LENGTH: 97 amino acids
1227                 (B)  TYPE: amino acid
1228                 (D)  TOPOLOGY: linear
1229
1230           (ii)  MOLECULE TYPE: protein
1231
1232           (xi)  SEQUENCE DESCRIPTION: SEQ ID NO:40:
1233
1234   Met Lys Ser Asn Asn Ala Leu Ile Val Ile Leu Gly Thr Val Thr Leu
1235    1               5              10                  15
1236
1237   Asp Ala Val Gly Ile Gly Leu Val Met Pro Val Leu Pro Gly Leu Leu
1238               20               25               30
1239
1240   Arg Asp Ile Val His Ser Asp Ser Ile Ala Ser His Tyr Gly Val Leu
1241           35               40               45
1242
1243   Leu Ala Leu Tyr Ala Leu Met Gln Phe Leu Cys Ala Pro Val Leu Gly
1244       50               55               60
1245
1246   Ala Leu Ser Asp Arg Phe Gly Arg Arg Pro Val Leu Leu Ala Ser Leu
1247    65               70               75                   80
1248
1249   Leu Gly Ala Thr Ile Asp Tyr Ala Ile Met Ala Thr Thr Pro Val Leu
1250               85               90               95
1251
1252   Trp
1253
1254
```

SEQUENCE VERIFICATION REPORT
PATENT APPLICATION *US/08/080,354*

INPUT SET: S6822.raw

Line	Error	Original Text
30	Wrong application Serial Number	(A) APPLICATION NUMBER: US
189	Entered (231) and Calc. Seq. Length (30) differ	(A) LENGTH: 231 base pairs
205	# of Sequences for line conflicts w/ running total	ATG GAG 48
210	# of Sequences for line conflicts w/ running total	CAG ATT GCC 96
215	# of Sequences for line conflicts w/ running total	GGT TAT GGT 144
220	# of Sequences for line conflicts w/ running total	TGT TGC 192
230	Sequence * thru 33 missing	(2) INFORMATION FOR SEQ ID NO:34:
256	Wrong Sequence Number	(2) INFORMATION FOR SEQ ID NO:10:
259	Entered (593) and Calc. Seq. Length (158) differ	(A) LENGTH: 593 base pairs
275	# of Sequences for line conflicts w/ running total	ATGAAAAATC TCATTGCTGA 60
278	# of Sequences for line conflicts w/ running total	GAACTGTGTG CGCAGGTAGA 120
281	# of Sequences for line conflicts w/ running total	GCGCAAAATG ACCAACAGCG 180
284	# of Sequences for line conflicts w/ running total	CCCGATGCCA GCATTCCTGA 240
287	# of Sequences for line conflicts w/ running total	TTGAAGCATC CTCGTCAGTA 300
290	# of Sequences for line conflicts w/ running total	GCCGAGACTT ATAGTCGCTT 360
293	# of Sequences for line conflicts w/ running total	AAAAGGGTAT CTAGAGGTTG 420
296	# of Sequences for line conflicts w/ running total	ATG TTC 469
301	# of Sequences for line conflicts w/ running total	GAG GAA 517
306	# of Sequences for line conflicts w/ running total	GCG ATA 565
316	Sequence 11 thru 36 missing	(2) INFORMATION FOR SEQ ID NO:37:
339	Wrong Sequence Number	(2) INFORMATION FOR SEQ ID NO:11:
342	Entered (1500) and Calc. Seq. Length (359) differ	(A) LENGTH: 1500 base pairs
358	# of Sequences for line conflicts w/ running total	ATGAAAAATC TCATTGCTGA 60
361	# of Sequences for line conflicts w/ running total	GAACTGTGTG CGCAGGTAGA 120
364	# of Sequences for line conflicts w/ running total	GCGCAAAATG ACCAACAGCG 180
367	# of Sequences for line conflicts w/ running total	CCCGATGCCA GCATTCCTGA 240
370	# of Sequences for line conflicts w/ running total	TTGAAGCATC CTCGTCAGTA 300
373	# of Sequences for line conflicts w/ running total	GCCGAGACTT ATAGTCGCTT 360
376	# of Sequences for line conflicts w/ running total	TCACGTAAAA AGGGTATCTA 420
379	# of Sequences for line conflicts w/ running total	CTT GCA TCT 470
384	# of Sequences for line conflicts w/ running total	ACT ACA 518
389	# of Sequences for line conflicts w/ running total	TTC AAG 566
394	# of Sequences for line conflicts w/ running total	ACT GTT 614
399	# of Sequences for line conflicts w/ running total	TTT TAC ACA 662
404	# of Sequences for line conflicts w/ running total	GAT GTG AAG 710
409	# of Sequences for line conflicts w/ running total	GTG GAG 758
414	# of Sequences for line conflicts w/ running total	GAG TTC 806
419	# of Sequences for line conflicts w/ running total	AGT TTT 854
424	# of Sequences for line conflicts w/ running total	GAA CGG ACT 902
429	# of Sequences for line conflicts w/ running total	TTT GGC 950
434	# of Sequences for line conflicts w/ running total	TCA GGA 998
439	# of Sequences for line conflicts w/ running total	GTG GAT 1046
444	# of Sequences for line conflicts w/ running total	TCC CGA 1094
449	# of Sequences for line conflicts w/ running total	ATG GGC CAG 1142
454	# of Sequences for line conflicts w/ running total	GTG GTA 1190
459	# of Sequences for line conflicts w/ running total	CAC TAAAATTCTC 1245
464	# of Sequences for line conflicts w/ running total	TTTATCACAG TTAAATTGCT 1305
467	# of Sequences for line conflicts w/ running total	CATCGTCATC CTCGGCACCG 1365

SEQUENCE VERIFICATION REPORT
PATENT APPLICATION *US/08/080,354*

INPUT SET: S6822.raw

Line	Error	Original Text
470	# of Sequences for line conflicts w/ running total	ACTGCCGGGC CTCTTGCGGG 1425
473	# of Sequences for line conflicts w/ running total	GCTCCTAGCG CTATATGCGT 1485
478	Sequence 12 thru 37 missing	(2) INFORMATION FOR SEQ ID NO:38:
541	Wrong Sequence Number	(2) INFORMATION FOR SEQ ID NO:12:
595	Entered (42) and Calc. Seq. Length (0) differ	(A) LENGTH: 42 base pairs
619	Entered (41) and Calc. Seq. Length (0) differ	(A) LENGTH: 41 base pairs
643	Entered (42) and Calc. Seq. Length (0) differ	(A) LENGTH: 42 base pairs
671	Sequence 18 missing	(2) INFORMATION FOR SEQ ID NO:19:
685	Wrong Sequence Number	(2) INFORMATION FOR SEQ ID NO:18:
688	Entered (41) and Calc. Seq. Length (0) differ	(A) LENGTH: 41 base pairs
709	Sequence 19 missing	(2) INFORMATION FOR SEQ ID NO:20:
712	Entered (64) and Calc. Seq. Length (24) differ	(A) LENGTH: 64 base pairs
735	# of Sequences for line conflicts w/ running total	AGG AAA CCC 49
744	Sequence 21 missing	(2) INFORMATION FOR SEQ ID NO:22:
761	Wrong Sequence Number	(2) INFORMATION FOR SEQ ID NO:21:
764	Entered (62) and Calc. Seq. Length (22) differ	(A) LENGTH: 62 base pairs
783	# of Sequences for line conflicts w/ running total	TCATGCCGCA AATGGCAATC 60
787	Sequence 22 missing	(2) INFORMATION FOR SEQ ID NO:23:
790	Entered (50) and Calc. Seq. Length (10) differ	(A) LENGTH: 50 base pairs
809	# of Sequences for line conflicts w/ running total	TTTGTTTGCG 50
814	Entered (50) and Calc. Seq. Length (10) differ	(A) LENGTH: 50 base pairs
833	# of Sequences for line conflicts w/ running total	CCGCAAGAGT 50
838	Entered (55) and Calc. Seq. Length (15) differ	(A) LENGTH: 55 base pairs
857	# of Sequences for line conflicts w/ running total	TGCAGTTTGT TTGCG 55
862	Entered (64) and Calc. Seq. Length (24) differ	(A) LENGTH: 64 base pairs
881	# of Sequences for line conflicts w/ running total	CCGCACAGAG TTTCGGGACC 60
888	Entered (84) and Calc. Seq. Length (44) differ	(A) LENGTH: 84 base pairs
907	# of Sequences for line conflicts w/ running total	GGTTGCCGTC GCAGCGGGCG 60
914	Entered (84) and Calc. Seq. Length (44) differ	(A) LENGTH: 84 base pairs
933	# of Sequences for line conflicts w/ running total	CGGCAACCGC CAGAGGAAGT 60
940	Entered (52) and Calc. Seq. Length (12) differ	(A) LENGTH: 52 base pairs
959	# of Sequences for line conflicts w/ running total	GTTGTACCAA AA 52
964	Entered (60) and Calc. Seq. Length (20) differ	(A) LENGTH: 60 base pairs
983	# of Sequences for line conflicts w/ running total	ATGCACTGTA GAGTTGTGCA 60
1037	Sequence 34 missing	(2) INFORMATION FOR SEQ ID NO:35:
1040	Entered (915) and Calc. Seq. Length (223) differ	(A) LENGTH: 915 base pairs
1056	# of Sequences for line conflicts w/ running total	ATG AGC 47
1061	# of Sequences for line conflicts w/ running total	CAG ACA CCT 95
1066	# of Sequences for line conflicts w/ running total	ACA GAA 143
1071	# of Sequences for line conflicts w/ running total	GAG CTG 191
1076	# of Sequences for line conflicts w/ running total	CAA CAA 239
1081	# of Sequences for line conflicts w/ running total	TTT AAG 287
1086	# of Sequences for line conflicts w/ running total	AGT CCT TCA 335
1091	# of Sequences for line conflicts w/ running total	AGA CAA 383
1096	# of Sequences for line conflicts w/ running total	ACC AAA 431
1101	# of Sequences for line conflicts w/ running total	ACATCTCGTA 482
1106	# of Sequences for line conflicts w/ running total	TCTATCAGGT CAATTCTCAT 542
1109	# of Sequences for line conflicts w/ running total	TATCACAGTT AAATTGCTAA 602

SEQUENCE VERIFICATION REPORT
PATENT APPLICATION *US/08/080,354*

INPUT SET: S6822.raw

Line	Error	Original Text
1112	# of Sequences for line conflicts w/ running total	TCGTCATCCT CGGCACCGTC 662
1115	# of Sequences for line conflicts w/ running total	TGCCGGGCCT CTTGCGGGAT 722
1118	# of Sequences for line conflicts w/ running total	TGCTAGCGCT ATATGCGTTG 782
1121	# of Sequences for line conflicts w/ running total	ACCGCTTTGG CCGCCGCCCA 842
1124	# of Sequences for line conflicts w/ running total	CGATCATGGC GACCACACCC 902
1170	Sequence 37 thru 38 missing	(2) INFORMATION FOR SEQ ID NO:39:
1173	Entered (301) and Calc. Seq. Length (49) differ	(A) LENGTH: 301 base pairs
1189	# of Sequences for line conflicts w/ running total	ACC GTC 48
1194	# of Sequences for line conflicts w/ running total	CCG GGC 96
1199	# of Sequences for line conflicts w/ running total	TAT GGC 144
1204	# of Sequences for line conflicts w/ running total	CCC GTT 192
1209	# of Sequences for line conflicts w/ running total	CTC GCT 240
1214	# of Sequences for line conflicts w/ running total	ACA CCC 288

SEQUENCE MISSING ITEM REPORT
PATENT APPLICATION   *US/08/080,354*

INPUT SET: S6822.raw

APPLICATION NUMBER
FILING DATE
PRIOR APPLICATION DATA

SEQUENCE CORRECTION REPORT
PATENT APPLICATION *US/08/080,354*

INPUT SET: S6822.raw

Line	Original Text	Corrected Text